



Article

On the Logic of a Prior Based Statistical Mechanics of Polydisperse Systems: The Case of Binary Mixtures

Fabien Paillusson 

School of Mathematics and Physics, University of Lincoln, Lincoln LN6 7TS, UK; fpaillusson@lincoln.ac.uk

Received: 30 May 2019; Accepted: 14 June 2019; Published: 16 June 2019



Abstract: Most undergraduate students who have followed a thermodynamics course would have been asked to evaluate the volume occupied by one mole of air under standard conditions of pressure and temperature. However, what is this task exactly referring to? If air is to be regarded as a mixture, under what circumstances can this mixture be considered as comprising only one component called “air” in classical statistical mechanics? Furthermore, following the paradigmatic Gibbs’ mixing thought experiment, if one mixes air from a container with air from another container, all other things being equal, should there be a change in entropy? The present paper addresses these questions by developing a prior-based statistical mechanics framework to characterise binary mixtures’ composition realisations and their effect on thermodynamic free energies and entropies. It is found that (a) there exist circumstances for which an ideal binary mixture is thermodynamically equivalent to a single component ideal gas and (b) even when mixing two substances identical in their underlying composition, entropy increase does occur for finite size systems. The nature of the contributions to this increase is then discussed.

Keywords: mixtures; entropy; polydispersity; binomial distribution

1. Introduction

Whether one thinks of ubiquitous substances such as air or drinking water or more specific fluids such as petroleum or milk, they all count as multicomponent systems, i.e., they all comprise more than one identifiable type of constituent. Furthermore, following early commentaries (see [1] and references therein) on Gibbs’ original work [2,3] on mixtures, the latter are thought to play a key role in the—quantum [4–13] or classical [14–25]—foundations of classical statistical mechanics; through the (in)famous *Gibbs paradox*. Somewhat surprisingly then, the vast majority of statistical mechanics textbooks covers principally single component systems. Mixtures are usually treated but as an exception to the identical particle paradigm. The reasons for this lack of visibility of multicomponent systems, despite their omnipresence in natural and artificial settings, are often rather elusive so that we could be left with—wrongly—attributing motives to their authors ranging from “mixtures do not matter as much as we think” to “mixtures are too complicated to treat in all their details anyways”. To be sure, if phase behaviour is to be considered in models of mixtures with interacting constituents, then characterising these phases and their occurrences is indeed a much more complicated problem than when looking at single component systems, as illustrated in relatively recent works [26,27]. Gibbs originally developed the notions of grand (and petit) canonical ensembles to elucidate the statistical thermodynamics of systems with varying particle numbers, including mixing problems [3]. More recently, the entropy of mixtures and Gibbs’ paradoxes were revisited within a more contemporary framework involving probabilities and particle exchanges protocols equivalent to the grand canonical ensemble for non-interacting systems [28]. However, for finite size systems, the grand canonical ensemble may not always give the same result as the canonical ensemble and, in practice, many mixing scenarios do not involve any external reservoir with which to exchange

particles. In addition, we will see that, even in absence of interaction between constituents, it is arguable that the very nature of a polydisperse system in the canonical ensemble requires additional care in its statistical mechanics treatment which impacts the finite size expected canonical free energy and entropy measures. The purpose of this paper is therefore two-fold: (1) to develop a mathematically and conceptually consistent statistical mechanics treatment of mixtures in the canonical ensemble and (2) to derive general finite size expressions for the canonical measures of free energy and entropy. From these expressions, we will then show that there are circumstances for which a mixture, in the canonical ensemble, becomes equivalent to a single component system thus partially justifying the apparent lack of coverage in the literature. In what follows, we will focus on the statistical mechanics of binary mixtures to illustrate the kind of difficulties that can already emerge at this rather simple level of polydispersity. The article is organised as follows: Section 2 reminds the reader of the classical textbook treatment of binary mixtures in the canonical ensemble. Section 3 introduces a heuristic generalisation of the textbook treatment developed in Section 2 and discusses some of its caveats. Section 4 proposes a general, prior based, statistical mechanics framework for binary mixtures and discusses its consequences. Section 5 looks at the problem of mixing between two mixtures with different compositions and identifies the various contributions to the entropy change so as provide a specific definition of mixing entropy. Finally, Section 6 discusses further perspectives and Section 7 presents conclusions.

2. Textbook Treatment of Binary Mixtures

The textbook treatment of binary mixtures of non-interacting particles with Hamiltonian H usually considers a system with a total of N particles with N_1 particles of type 1 and mass m_1 and $N_2 = N - N_1$ particles of type 2 and mass m_2 , different from type 1, confined in a box of volume V and maintained at a temperature $T = (k_B \beta)^{-1}$. Following the $1/n!$ prescription for any n identical particles in the system, the classical canonical partition function $Q(N, N_1, \beta, V)$ of the system then reads (e.g., in Refs. [8,9,11]):

$$Q(N, N_1, \beta, V) = \frac{1}{N_1!(N - N_1)!} \int_{\text{phase space}} \frac{\prod_{i=1}^N d^3 r_i d^3 p_i}{h^{3N}} e^{-\beta H}, \quad (1)$$

where h —usually taken as the Planck constant—is a quantity with the dimension of an action and the Hamiltonian H reads

$$H = U_{\text{box}}(\{\vec{r}_i\}_{i=1,\dots,N}) + \sum_{i=1}^{N_1} \frac{\vec{p}_i^2}{2m_1} + \sum_{i=1}^{N-N_1} \frac{\vec{p}_i^2}{2m_2}, \quad (2)$$

where $U_{\text{box}}(\{\vec{r}_i\}_{i=1,\dots,N})$ is infinite if any particle goes outside the bounding box and is zero otherwise. Upon choosing, as is often done, $N_1 = N/2$, it follows that:

$$Q(N, N_1 = N/2, \beta, V) = \frac{1}{N_1!(N - N_1)!} \frac{V^N}{\Lambda_1^{3N_1} \Lambda_2^{3(N-N_1)}} = \frac{1}{(N/2)!^2} \frac{V^N}{\Lambda_1^{3N/2} \Lambda_2^{3N/2}}, \quad (3)$$

where $\Lambda_i \equiv h\sqrt{\beta/(2\pi m_i)}$ is the thermal wavelength of species i . Finally, applying the Stirling approximation $\ln N! \approx N \ln N - N$, one finds for sufficiently large N the free energy

$$\beta F(N, N_1 = N/2, V, \beta) \simeq -N \ln 2 - N + \frac{N}{2} \ln(\rho \Lambda_1^3) + \frac{N}{2} \ln(\rho \Lambda_2^3), \quad (4)$$

where $\rho \equiv N/V$. In Equation (4), the second and third terms can be interpreted as the free energy that a gas of $N/2$ identical particles of respectively type 1 and 2 would have, had they been separated in equal sized compartments of volume $V/2$. Given that $F = U - TS$ and that T does not vary, it is then rather timely to introduce the mixing entropy ΔS_{mix} as being:

$$\frac{\Delta S_{\text{mix}}}{k_B} \equiv \beta F(N_1 = N/2, \beta, V/2) + \beta F(N_2 = N/2, \beta, V/2) - \beta F(N, N_1 = N/2, \beta, V) = N \ln 2. \quad (5)$$

The entropy gain of $k_B \ln 2$ per particle found in Equation (5) upon mixing is the well known entropy of mixing attributable to the difference in identity between the two species 1 and 2. Technically speaking, this difference in identity is betrayed by the factor $(N/2!)^{-2}$ in Equation (3), instead of the usual $1/N!$ expected for identical particles. However, it has been argued that interpreting the $Nk_B \ln 2$ entropy change as stemming from this combinatorial factors was in fact misleading [11]. We will come back to this issue in Section 5.

3. Heuristic Generalisation

One of the issues with what was presented above as the “textbook” derivation is that it may appear somewhat artificial in a general context. For example, in the case where $N_1 = N/2$, it follows that N ought to be an even number as there cannot be a fractional number of particles. In addition, even in absence of fractional particle numbers, how are we to choose the species upon adding a single additional particle to the overall system? These questions become more and more inevitable as the kind of mixtures one wants to look at becomes more complex with, e.g., $N_1 = N/\sqrt{2}$ or when considering more than two species or even infinitely many.

One possible way of tackling these issues is to interpret the composition of a mixture through the probability $p \in [0, 1]$ for a particle in the box to be of type 1 and then recast N_1 as $N_1 \equiv Np$. Since p is a probability, the product Np need not be an integer and each new particle added to the mixture will be of type 1 with probability p and type 2 with probability $1 - p$. With this new prescription, Equation (3) can be rewritten as follows:

$$Q(N, p, \beta, V) = \frac{1}{\Gamma(Np + 1)\Gamma(N(1 - p) + 1)} \frac{V^N}{\Lambda_1^{3Np}\Lambda_2^{3N(1-p)}}, \quad (6)$$

where $\Gamma(x + 1) \equiv \int_0^{+\infty} dy y^x e^{-y}$ is the Euler gamma function which generalises the factorial function to the reals. For large values of x , one can use a saddle point approximation and finds that $\Gamma(x + 1)$ satisfies the Stirling approximation $\Gamma(x + 1) \simeq x \ln x - x$. Thus, in the large N limit, one finds for the free energy

$$\beta F(N, p, \beta, V) \simeq Np \ln(\rho \Lambda_1^3) + N(1 - p) \ln(\rho \Lambda_2^3) - N - Ns(p), \quad (7)$$

where $s(p) \equiv -p \ln p - (1 - p) \ln(1 - p)$. Although some authors [19,26] refer to $s(p)$ as the mixing entropy, we follow [29] and consider that the mixing entropy denomination should be left for unambiguously prepared mixing scenarios and their corresponding entropy variations and therefore refer to $s(p)$, appearing in the equilibrium free energy of a mixture, as its *composition entropy*. For the reader who is more mathematically inclined, in the case of a Bernoulli random variable with probability p , $s(p)$ is also called the *binary entropy* [30].

It can be noted that Equation (7) can be rewritten, without any explicit reference to different species, as follows:

$$\beta F(N, p, \beta, V) \simeq N \left(\ln(\rho \tilde{\Lambda}^3) - 1 \right) - Ns(p), \quad (8)$$

where $\tilde{\Lambda} \equiv h\sqrt{\beta/(2\pi\tilde{m})}$ and $\tilde{m} \equiv m_1^p m_2^{1-p}$ is the weighted geometric mean of the masses of the species 1 and 2. One important consequence of Equation (8) is that, at *fixed* composition, both $\tilde{\Lambda}$ and $s(p)$ have fixed values and do not contribute to the thermodynamic properties of the system and therefore the non-interacting binary mixture is thermodynamically equivalent to a system of N effective identical particles [22].

Suppose now that we have two binary mixtures each comprising species 1 and 2 but with different compositions: mixture A with probability p_A for a particle of the mixture to be of type 1 and mixture B with probability p_B for a particle of the mixture to be of type 2. Each mixture contains $N/2$ particles and is initially confined in a box of volume $V/2$ (cf. Figure 1).

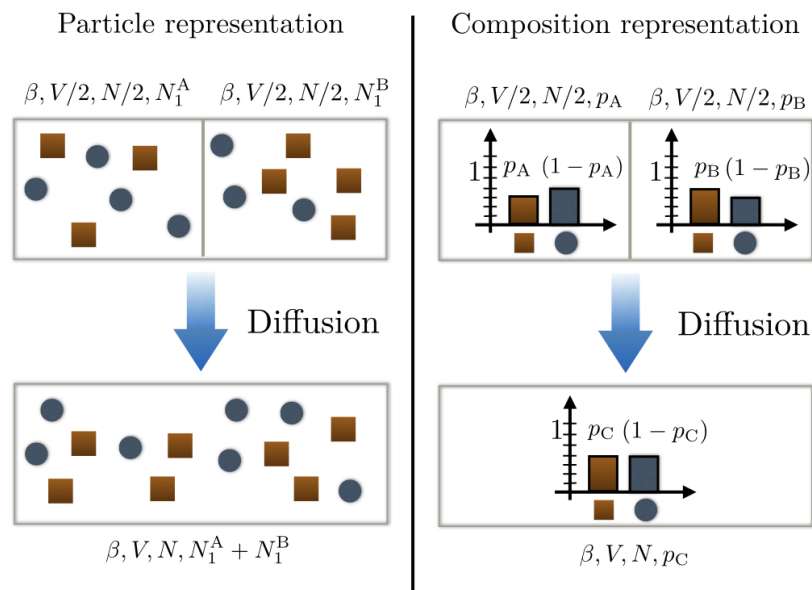


Figure 1. Schematic representation of a mixing scenario between two different substances. The top left panel uses a particle representation with type 1 particles as dark orange squares and type 2 particles as blue disks. We see that the compositions of substances A and B are different. This is further illustrated in representing their underlying probability distribution on the top right panel. Upon mixing the bottom panels, they form a new composition C that is a priori different from A and B.

They are then mixed together in a box of volume V : what is the entropy change in the mixture? To address this question, we need first to determine the composition of the new mixture C obtained after mixing. In the absence of chemical reactions, the conservation of particle number compels us to ascribe the probability $p_C = (p_A + p_B)/2$. The entropy change upon mixing reads then

$$\frac{\Delta S_{mix}}{k_B} = \beta F(N/2, p_A, \beta, V/2) + \beta F(N/2, p_B, \beta, V/2) - \beta F(N, p_C, \beta, V) = D_{JS}(p_A|p_B), \quad (9)$$

where

$$D_{JS}(p_A|p_B) = \frac{1}{2} \sum_{i=A,B} \left[p_i \ln \left(\frac{2p_i}{p_A + p_B} \right) + (1 - p_i) \ln \left(\frac{2(1 - p_i)}{(1 - p_A) + (1 - p_B)} \right) \right]. \quad (10)$$

is the *Jensen–Shannon divergence*. Two features of D_{JS} worth mentioning is that it is positive definite and bounded from above by $\ln 2$ and its square root is a distance between probability distributions [31]. One can verify for example that if $p_A = 0$ and $p_B = 1$, then $D_{JS} = \ln 2$ thus retrieving Equation (5).

In spite of the generalisation of the textbook derivation to any binary mixture and the physical insights gained from adopting a *composition as probability* paradigm, the approach developed above suffers from a few mathematical and conceptual problems which need to be addressed:

- The heuristic approach used in this section starts off by directly generalising Equation (3) into Equation (6). Mathematically, it cannot start from Equation (1) as, in general, it could involve a fractional number of phase space integrals over particles of type 1 or 2, respectively. Thus, the current approach cannot be used as a basis to devise a mathematically rigorous composition-probability-based statistical mechanics of mixtures.
- Equation (6) made use of the Euler Gamma function to replace the more traditionally accepted $n!$ terms at the denominator so as to account for Np not being an integer. While mathematically this may be fine, it is not justified within statistical mechanics itself and, indeed, as the first point was being raised, it is hardly so, as one could have a fractional number of phase space integrals.

- Problems arise with Np as well. What does it mean to switch from N_1 (an integer) to Np (not necessarily an integer) in the canonical ensemble? If p is a probability for a particle to be of type 1, then the particle type is a Bernoulli random variable $\mathbf{t} = \{1, 2\}$ such that $p(\mathbf{t} = 1) = p$ and $p(\mathbf{t} = 2) = 1 - p$. Suppose we now model a mixture as a collection $\{\mathbf{t}_i\}_{i=1,\dots,N}$ of N independent such random variables. This leads us to define the random variable $\mathbf{N}_1 \equiv \sum_{i=1}^N (-\mathbf{t}_i + 2)$ corresponding to the number of particles of type 1 in the system. If we denote $\langle\langle\cdot\rangle\rangle$, from the composition average, $\langle\langle\mathbf{N}_1\rangle\rangle = Np$ comes. Thus, by substituting N_1 with Np in Equation (3), one effectively replaces an integral number of particles by a real positive expectation value. Conceptually, this poses a problem, at least in principle, as any given individual mixture will only ever have an integer number of particles usually close to, but different from, Np .

It is the aim of the following sections to devise a rigorous statistical mechanics framework of binary mixtures for which the composition is interpreted as an a priori probability distribution with probability p for any given particle to be of type 1 (and probability $(1 - p)$ for a particle to be of type 2). Only once such a framework has been laid out can we hope to delineate the range of validity, if any, of the results summarised in Equations (7)–(9) and the approximate reasoning used to derive them.

4. Prior Based Statistical Mechanics of a Binary Mixture

We now consider the simple case of the composition of a binary mixture modelled as a collection $\{\mathbf{t}_i\}_{i=1,\dots,N}$ of N independent Bernoulli random variables with $p(\mathbf{t}_i = 1) = p$ and $p(\mathbf{t}_i = 2) = 1 - p$. Denoting \mathbf{N}_1 as the random variable for the number of particles of type 1 among N , the probability $P(\mathbf{N}_1 = N_1 | N, p)$ for it to be equal to a set integer value N_1 follows a Binomial distribution $B_{N,p}(N_1)$ defined by:

$$P(\mathbf{N}_1 = N_1 | N, p) = B_{N,p}(N_1) \equiv \frac{N!}{N_1!(N - N_1)!} p^{N_1} (1 - p)^{N - N_1}. \quad (11)$$

Given the status of random variable of \mathbf{N}_1 , it means that the set of values it can take corresponds to a set of *different* possible realisations of the *same* mixture. As a side note, in models with interacting components, the fact that there can be multiple realisations of the same mixture (as characterised by an a priori probability distribution) translates into using random matrices to set the interaction strengths between constituents [27].

From Equation (11), we see that

$$\frac{1}{N_1!(N - N_1)!} = \frac{B_{N,p}(N_1)}{N! p^{N_1} (1 - p)^{N - N_1}}. \quad (12)$$

Now, for any given mixture realisation, both N and N_1 are fixed and the framework of the canonical ensemble applies, including Equation (3). We can then substitute $1/(N_1!(N - N_1)!)$ by its expression in Equation (12) and get:

$$Q(N, \mathbf{N}_1 = N_1, \beta, V) = \frac{V^N}{\Lambda_1^{3N_1} \Lambda_2^{3(N - N_1)}} \frac{B_{N,p}(N_1)}{N! p^{N_1} (1 - p)^{N - N_1}}. \quad (13)$$

We seek a definition of the free energy of a mixture that would be realisation-independent. This is because repeating an experiment with a given mixture—characterised by prior composition probability—likely involves different realisations of its composition. To this end, we denote $\mathcal{F}(N, p, V, \beta) \equiv \langle\langle -k_B T \ln Q(N, \mathbf{N}_1, \beta, V) \rangle\rangle$ the canonical free energy averaged over realisations of \mathbf{N}_1 . After a bit of algebra and using the fact that $\langle\langle\mathbf{N}_1\rangle\rangle = Np$, we get

$$\beta \mathcal{F}(N, p, V, \beta) = -N \ln V + N \ln \tilde{\Lambda}^3 + \ln N! - Ns(p) + H(B_{N,p}, B_{N,p}), \quad (14)$$

where $s(p)$ and $\tilde{\Lambda}$ are as introduced in Equations (7) and (8), respectively, and $H(f, g) \equiv -\sum_{N_1=0}^N f(N_1) \ln g(N_1)$ making then $H(B_{N,p}, B_{N,p})$ interpretable as the *realisation entropy* of the

mixture. There does not exist any exact closed form expression for $H(B_{N,p}|B_{N,p})$ [32] but for N sufficiently large $H(B_{N,p}, B_{N,p}) \sim (1/2) \ln(2\pi e N p(1-p))$ (see Appendix A) so that, in the large N limit, Equation (14) is equivalent to Equation (8), thus justifying more rigorously its validity.

A few remarks are in order regarding Equation (14):

- The large N limit leading to Equation (8) in the heuristic derivation (Section 3) implies that both Np and $N(1-p)$ should be sufficiently large for the Stirling approximation to hold, which can require a very large N if p is either very small or very close to unity. On the contrary, Equation (14) converges rather quickly with N to the asymptotic form Equation (8) *especially* when p is either very small or very close to unity. This can be explained by remarking that, if $1-p$ is very small and N finite, any realisation will most likely have very few, if any, particles of type 2. As a consequence, the realisation entropy will be closer to zero in magnitude as there is less uncertainty in the composition realisations. Therefore, although the heuristic derivation leads to the asymptotic form Equation (8), it incorrectly—given that Equation (14) has stronger mathematical and conceptual foundations—predicts the composition for which the asymptotic regime is reached the fastest.
- For finite N , the realisation entropy term in Equation (14) actually *decreases* the entropy estimate given by the first four terms in the r.h.s of Equation (14) as it contributes *negatively* to the entropy of the system. This can be understood by adopting a “surprise” interpretation of entropy. The last two terms of Equation (14) can then be interpreted as the average surprise to have N_1 type 1 particles in the mixture. On the one hand, the $Ns(p)$ contribution to entropy stems from an estimation of the surprise for a given realisation N_1 as being $N_1 \ln p + (N - N_1) \ln(1-p)$. This would be exact if one were to either assign a particular order to the particles or if one were to repeat single-particles experiments N times, where the identity of the particle for each try is obtained from the underlying probability distribution of the Bernoulli variable \mathbf{t} , and add-up the observed individual surprises [22]. On another hand, the probability to have N_1 particles of type 1 among N is $B_{N,p}(N_1)$ and the corresponding surprise is $-\ln B_{N,p}(N_1)$. The difference between the two gives us the *relative surprise* $-\ln \frac{p^{N_1}(1-p)^{N-N_1}}{B_{N,p}(N_1)}$ which captures the contribution to entropy owing to composition.

Finally, we note that Sollich et al. [26,33] have proposed a strategy based on a moment-description of thermodynamic quantities for sufficiently large N to tackle the fact that a mixture composition of a single finite system is but a realisation of some underlying prior composition. This strategy differs in spirit from the one developed in the present paper in that it aims at performing a dimensional reduction of the free energy landscape by using a dependence of the free energy in the moments (ideally a small number of them) of the feature used to characterise the composition (e.g., size, mass, etc...) to obtain insights on the phase behaviour of mixtures. This objective is currently beyond the scope of the present paper.

5. Identifying the Gibbs Mixing Entropy

In this section, we consider a situation analogous to the one described in Figure 1 whereby $N/2$ particles (N being even) of a mixture A initially confined in a volume $V/2$ mix with $N/2$ particles of a mixture B initially confined in a volume $V/2$. Mixtures A and B only comprise type 1 and 2 particles but each particle identity is modelled with different random variables depending on the mixture it belongs to: \mathbf{t}^A with $p(\mathbf{t}^A = 1) = p_A$ and \mathbf{t}^B with $p(\mathbf{t}^B = 1) = p_B$. Following the theory developed in Section 4, we model the composition of mixture A (resp. B) as a collection of $N/2$ independent Bernoulli random variables $\{\mathbf{t}_i^A\}_{i=1,\dots,N/2}$ (resp. $\{\mathbf{t}_i^B\}_{i=1,\dots,N/2}$) and denote \mathbf{N}_1^A (resp. \mathbf{N}_1^B) the random variable giving the number of type 1 particles in mixture A (resp. mixture B). \mathbf{N}_1^A and \mathbf{N}_1^B both follow

a Binomial distribution and therefore, prior to mixing, the free energy of each mixture—for a given realisation—follows from Equation (12):

$$\beta F_{A/B}(N/2, \mathbf{N}_1^{A/B} = N_1^{A/B}, \beta, V/2) = -\ln \left[\frac{V^{N/2}}{\Lambda_1^{3N_1^{A/B}} \Lambda_2^{3(N/2 - N_1^{A/B})}} \frac{B_{N/2, p_{A/B}}(N_1^{A/B})}{\frac{N}{2}! p_{A/B}^{N_1^{A/B}} (1 - p_{A/B})^{N/2 - N_1^{A/B}}} \right]. \quad (15)$$

Upon mixing, the A and B mixtures will form a new mixture C with a composition emerging from the constraints during the diffusion process [29]. In the absence of chemical reactions, the number \mathbf{N}_1^C of particles of type 1, once mixing has occurred, is a random variable satisfying:

$$\mathbf{N}_1^C \equiv \mathbf{N}_1^A + \mathbf{N}_1^B, \quad (16)$$

and

$$\langle \langle \mathbf{N}_1^C \rangle \rangle = \frac{N}{2}(p_A + p_B) = Np_C. \quad (17)$$

We note that Equation (17) justifies the approach used for mixture C within the heuristic approach described in Section 3. If we want to know what is the probability for having $\mathbf{N}_1^C = N_1^C$, then we have:

$$P(\mathbf{N}_1^C = N_1^C | N, p_A, p_B) = P_{N, p_A, p_B}(\mathbf{N}_1^C) \equiv \sum_{N_1^A=0}^{N/2} \sum_{N_1^B=0}^{N/2} B_{N/2, p_A}(N_1^A) B_{N/2, p_B}(N_1^B) \delta_{N_1^C, N_1^A + N_1^B} \quad (18)$$

$$= \sum_{N_1^A=0}^{N_1^C} B_{N/2, p_A}(N_1^A) B_{N/2, p_B}(N_1^C - N_1^A), \quad (19)$$

where $\delta_{i,j} = 1$ if $i = j$, and zero otherwise, is the Kronecker delta function. If $p_A = p_B$, then $P_{N, p_A, p_B}(\mathbf{N}_1^C) = B_{N, p_A}(N_1^C)$. However, there is no known closed form expression for $P_{N, p_A, p_B}(\mathbf{N}_1^C)$ when $p_A \neq p_B$ so if one wants to estimate its values it has to be done either numerically by carrying out the whole summation or by using an approximate expression [30].

The canonical free energy of the final state of the system for a given realisation \mathbf{N}_1^C can be written from the form of the canonical partition function in Equation (13):

$$F_C(N, \mathbf{N}_1^C = N_1^C, \beta, V) = -\ln \left[\frac{V^N}{\Lambda_1^{3N_1^C} \Lambda_2^{3(N - N_1^C)}} \frac{B_{N, p_C}(N_1^C)}{N! p_C^{N_1^C} (1 - p_C)^{N - N_1^C}} \right]. \quad (20)$$

The difference with Equation (13), however, is that, in Equation (20), the probability distribution $B_{N, p_C}(N_1^C)$ is in general not equal to the actual probability distribution (given in Equation (19)) for obtaining N_1^C type 1 particles after the mixing of substances A and B. We will see in what follows that, as long as $B_{N, p_C}(N_1^C)$ has the same support as $P(\mathbf{N}_1^C = N_1^C | N, p_A, p_B)$, this problem can be overcome.

For a given realisation of mixtures A and B composition, we can now obtain the entropy variation ΔS_{mix} upon mixing as being of the two gases from $\Delta S_{mix}/k_B = \beta F_A(N/2, \mathbf{N}_1^A, \beta, V) + \beta F_B(N/2, \mathbf{N}_1^B, \beta, V) - \beta F_C(N, \mathbf{N}_1^C, \beta, V)$. For each realisation, a different entropy variation can be found in principle. Like in Section 4, a realisation-free entropy change $\Delta S_{mix} \equiv \langle \langle \Delta S_{mix} \rangle \rangle$ can be sought by averaging over composition realisations of substances A and B. We get (see Appendix B):

$$\frac{\Delta S_{mix}}{k_B} = \underbrace{\ln \left(\frac{2^N ((N/2)!)^2}{N!} \right)}_{\text{partitioning entropy}} + \underbrace{N D_{JS}(p_A | p_B)}_{\text{square metric}} + \underbrace{\sum_{i=A,B} H(B_{N/2, p_i}) - H(P_{N, p_A, p_B})}_{\text{realisation entropy}}. \quad (21)$$

In Equation (21), we have identified three different contributions to the entropy change upon mixing substances A and B that are worth commenting on:

- **Partitioning:** The first term on the r.h.s of Equation (21) corresponds to the entropy change owing to having removed a partition between the initial compartments or, said differently, results from the more numerous partitioning possibilities offered when the whole volume V can be explored instead of $V/2$. Indeed, when each particle can freely roam in the box of volume V , imagining a virtual division splitting the box into two equal halves will lead each particle to be either on one side or the other of the virtual dividing wall. There are then 2^N possibilities. When the dividing wall switches from virtual to real, and $N/2$ particles have to be on each side, the number of ways of realising this partitioning is $N!/((N/2)!)^2$. The corresponding entropy change is therefore larger than zero for any finite N .
- **Composition distance:** The second term on the r.h.s of Equation (21) has already been discussed in Section 4 and corresponds to a contribution to the entropy variation stemming from how different are the compositions of mixtures A and B when characterised by underlying probability distributions for particle identities. Since $D_{JS}(p_A|p_B)$ is a square metric, it becomes strictly zero when the a priori compositions are identical.
- **Composition realisation:** The third and last term identified on the r.h.s of Equation (21) corresponds to the entropy variation owing to realisation considerations. It is worth noting that when the compositions are identical i.e., $p_A = p_B$ then $P_{N,p_A,p_B} = B_{N,p_A}$ but $H(B_{N,p_A}, B_{N,p_A}) \leq 2H(B_{\frac{N}{2},p_A}, B_{\frac{N}{2},p_A})$ because of the *submodular* character of Shannon's entropy function [34]. As a result, the entropy variation due to composition realisations is larger than or equal to zero upon mixing even when substances have the same underlying composition.

When split into the three contributions identified in Equation (21), the reasons for an entropy increase upon mixing can be seen in a new light.

First, as noted in [11], conflating the presence of combinatorial terms in an entropy expression with an explicit role of particle identity can be misleading. Indeed, the partitioning entropy term has nothing to do with whether substances A and B are considered identical or not and is therefore not exclusive to the mixing of different substances. It will therefore appear in all mixing scenarios no matter what. This positive increase in entropy owing to partitioning can, at least in principle, be used to do meaningful work as exemplified with Szilard's engine [35]. This positive increase in entropy also further supports the—usually dismissed—intuitive claim that, even when substances are identical, some form of mixing does indeed occur.

The realisation entropy contribution that is positive even when substances have an identical underlying composition can be understood by the fact that upon mixing information on the initial realisations of mixtures A and B has disappeared. Therefore, if one were to insert a dividing wall, even if it happens that there are $N/2$ particles on each side, the realised compositions on either sides of the wall will be differing from the initial ones, even if selective membrane is used. This difference in realisation can in principle be used to do work by adapting Szilard's engine to incorporate a semi-permeable membrane. This conclusion would not follow, however, if the substances are identical in their underlying composition and p_A is exactly equal to 0 or 1 since the consideration on realisations would then become meaningless.

It remains then the square metric contribution to entropy which is the only one to exactly vanish if substances are identical in their underlying composition and to measure in what sense substances A and B differ from each other in principle. For reasons detailed elsewhere [22], we argue that this contribution reflects best Gibbs' original insights on the mixing of substances by diffusion [2] and shall therefore denote $\Delta S_{mix}^{Gibbs} \equiv k_B N D_{JS}(p_A|p_B)$.

All of the above considerations with regard to Equation (21) hold for finite size systems. From a standard use of Stirling's approximation, the first term on the r.h.s of Equation (21) vanishes in the thermodynamic limit. At first glance, it is nevertheless unclear how the $H(P_{N,p_A,p_B}, B_{N,p_C})$ behaves in the large N limit. It can be shown (Appendix A) that the term $H(P_{N,p_A,p_B}, B_{N,p_C})$ is at most of order $\sim \mathcal{O}(\ln N)$ so that

$$\Delta S_{mix} \simeq \Delta S_{mix}^{Gibbs} = k_B N D_{JS}(p_A|p_B) \quad (22)$$

for large N .

6. Discussion

After having briefly presented what we called the textbook treatment of binary mixtures, we looked at a heuristic generalisation which is grounded in the idea that a size-independent definition of a substance necessitates the existence of a prior underlying probability distribution from which are drawn the particles identity. The heuristic formulation works by substituting N_1 the number of particles of type 1 by Np , p being the probability for an individual particle to be of type 1. Following some postulated extensions of the canonical ensemble framework, this led us to Equation (13) which enabled us to discuss how the thermodynamics of a binary mixture can then be equivalent to that of a system of identical particles, provided the composition is kept fixed. Acknowledging various caveats of the heuristic approach, we then developed a more rigorous approach fully compatible with traditional statistical mechanics. This new approach starts by taking seriously the idea of a prior probability distribution underlying the composition of a substance. If that is the case, then two different realisations of the same underlying distribution need not have the exact same composition. Having this in mind, it then conjectured that to obtain a realisation-independent thermodynamics of a binary mixture, one needs to consider the free energies averaged over realisations. This procedure enabled us to retrieve Equation (13) in the large N limit, thus validating the more intuitive approach used in Section 3. Incidentally, this further supports the idea that for dilute gases or substances at fixed composition, the actual polydisperse character of the substance bears no consequences on the thermodynamics of the system and it is enough to consider the system as comprising only—effective—identical particles. It is unclear whether this is purposeful or not, but this equivalence between a binary mixture and a single component system can partially justify—at least a posteriori—the predominance of single component systems in the literature available at undergraduate level.

Aside from the average free energy of a given mixture, the entropy variation upon mixing two different mixtures has also been studied within the new statistical mechanics framework proposed in this paper. It was found that the entropy of mixing comprised three physically different contributions: one owing to partitioning that is not exclusive to the mixing of different substances, one owing to a square distance between the underlying probability distributions for particles' identity and a last one owing to realisation entropy differences. This last term is new and does not vanish when the underlying compositions of the two mixed substances are the same. Upon inspection, the second—metric based—term in Equation (21) is therefore the closest to the entropy of mixing, as discussed by Gibbs in [2]. Finally, we discussed the thermodynamic limit behaviour of Equation (21) and showed that it reduced to the Gibbs entropy (22).

It is worth noting that the proposed theory of binary mixtures relies on the assumption that the substances can be modelled as collections of independent random—identity—variables. In practice, any protocol which behaves the same way with any particle regardless of how many of a specific kind it has already processed, would give rise to a Binomial distribution for particle number of type 1 (resp. 2). However, there can also be non-independent collections of Bernoulli variables with $P_{N,p}(N_1) \neq B_{N,p}(N_1)$ (e.g., kinetic proof reading or active sorting) [36]. In such cases, the mathematical framework developed in these pages is still valid, but the asymptotic behaviours of average free energies need to be elucidated for each individual case. In addition, it is important to stress that, if the particles were to interact with each other, that would not necessarily mean that the particles' identity random variables are not independent. Instead, particle interactions could simply be expressed through random matrices [27]. Therefore, one avenue of exploration would be to adapt it to interacting mixtures. Moreover, it is possible to see a parallel with the composition-realisation-based statistical mechanics developed in the present paper and averages over disorder in systems with quenched disorder (see e.g., [37]). This offers possibilities of cross-fertilisation between different branches of statistical mechanics. Now, if the realisation-based approach corresponds to averages over quenched disorder in other fields, it is tempting to draw a parallel between grand canonical

ensemble-treatments of mixtures [28] and annealed disorder. The author is not aware of an equation equivalent to Equation (22) in the context of the grand canonical ensemble and this is another route worth exploring.

Finally, in Refs. [22,29], the heuristic approach was used to suggest that both discrete and continuous polydisperse systems were equivalent to a system of identical particles, provided the composition remains the same and a generalised version of Equation (9) was derived for mixtures of arbitrary number of components and composition. Whether these results can remain valid in a version of the present framework extended to more than two components remains to be determined.

7. Conclusions

In this article we proposed a theory to address the possibility of different realisations of a given composition and obtain realisation-independent free energies in the canonical ensemble. This lead us to find that (a) in addition to the composition entropy discussed in [22,29] a realisation entropy—vanishing in the thermodynamic limit—emerges in the expression of the free energy of a binary mixture, (b) for a fixed composition the thermodynamics of a binary mixture is equivalent to that of a system of identical particles, (c) the entropy change upon mixing two finite size identical substances has two positive contributions owing to firstly an increase in partitioning entropy and secondly a loss of information on substances' realisations and (d) in the thermodynamic limit the mixing entropy is a square norm between the compositions of the substances being mixed. Further work remains to be done to extent this work to more complex mixtures.

Funding: This research received no external funding.

Conflicts of Interest: The author declares no conflict of interest.

Appendix A. Limiting Behaviour of $H(P_{N,p_A,p_B}, B_{N,p_C})$

We wish to determine the asymptotic behaviour of $H(P_{N,p_A,p_B}, B_{N,p_C})$ as N tends to infinity. To this end, we first note that $H(P_{N,p_A,p_B}, B_{N,p_C}) \geq 0$ as it is the cross entropy over a discrete probability distribution. Next, we recall Jensen's inequality [38], which, for a given random variable \mathbf{X} , states that $E[\varphi(\mathbf{X})] \geq \varphi(E[\mathbf{X}])$, where $E[\cdot]$ stands for the expectation value. If we choose $B_{N,p_C}(\mathbf{N}_1^C)/P_{N,p_A,p_B}(\mathbf{N}_1^C)$ for \mathbf{X} and $\varphi(x) = \ln x$, we get

$$E[\varphi(\mathbf{X})] = \sum_{N_1^C=0}^N P_{N,p_A,p_B}(N_1^C) \ln \left(\frac{B_{N,p_C}(N_1^C)}{P_{N,p_A,p_B}(N_1^C)} \right) \geq \ln \left[\sum_{N_1^C=0}^N B_{N,p_C} \right] = 0, \quad (\text{A1})$$

which yields

$$H(P_{N,p_A,p_B}, B_{N,p_C}) \leq H(P_{N,p_A,p_B}, P_{N,p_A,p_B}). \quad (\text{A2})$$

Note that Equation (A1) is valid because B_{N,p_C} and P_{N,p_A,p_B} have the same support.

Next, we seek to evaluate the limiting behaviour of $H(P_{N,p_A,p_B}, P_{N,p_A,p_B})$. We recall that P_{N,p_A,p_B} stands for the probability distribution of the sum of two binomial variables distributed according to $B_{\frac{N}{2},p_A}$ and $B_{\frac{N}{2},p_B}$, respectively. In the large N limit, the central limit theorem applies and we have

$$B_{\frac{N}{2},p_{A/B}}(N_1^{A/B}) \sim \mathcal{N}_{\frac{Np_{A/B}}{2}, \frac{Np_{A/B}(1-p_{A/B})}{2}}(N_1^{A/B}), \quad (\text{A3})$$

where

$$\mathcal{N}_{\mu,\sigma^2}(x) \equiv \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \quad (\text{A4})$$

Given that $\mathbf{N}_1^C = \mathbf{N}_1^A + \mathbf{N}_1^B$, we can use Equation (A3) and consider \mathbf{N}_1^C as a sum of two normally distributed random variables. Since the sum of two normally distributed random variables with mean

μ_A (resp. μ_B) and variance σ_A^2 (resp. σ_B^2) is also normally distributed but with mean $\mu_A + \mu_B$ and variance $\sigma_A^2 + \sigma_B^2$,

$$P_{N,p_A,p_B}(N_1^C) \sim \mathcal{N}_{N_{p_C}, \frac{N}{2}(p_A(1-p_A)+p_B(1-p_B))}(N_1^C). \quad (A5)$$

is provided.

It is straightforward to show that

$$H(\mathcal{N}_{\mu,\sigma^2}, \mathcal{N}_{\mu,\sigma^2}) = \frac{1}{2} \ln(2\pi e\sigma^2). \quad (A6)$$

Thus, denoting $\sigma_C^2 = N(p_A(1-p_A) + p_B(1-p_B))/2$, we get that

$$H(P_{N,p_A,p_B}, P_{N,p_A,p_B}) \sim H(\mathcal{N}_{N_{p_C}, \sigma_C^2}, \mathcal{N}_{N_{p_C}, \sigma_C^2}) \sim \mathcal{O}(\ln N), \quad (A7)$$

which, when combined with Equation (A2), shows that $H(P_{N,p_A,p_B}, B_{N,p_C})$ is at most of order $\sim \mathcal{O}(\ln N)$ in the large N limit.

Appendix B. Deriving the Entropy Change Averaged over Compositions

In this section, we want to find an expression for the entropy variation upon mixing two substances averages over all possible realisations of their composition. We start by discussing more formally which average is taken. In our particular case of mixing, the average of any function $f(N_1^A, N_1^B)$ is defined as

$$\langle \langle f(N_1^A, N_1^B) \rangle \rangle \equiv \sum_{N_1^A=0}^{N/2} \sum_{N_1^B=0}^{N/2} B_{\frac{N}{2}, p_A}(N_1^A) B_{\frac{N}{2}, p_B}(N_1^B) f(N_1^A, N_1^B). \quad (A8)$$

From the general definition above, two specific cases are worth looking at:

- If $f(N_1^A, N_1^B) = f(N_1^{A/B})$ i.e., only depends on one of the variables, then

$$\langle \langle f(N_1^A, N_1^B) \rangle \rangle = \sum_{N_1^{A/B}=0}^{N/2} B_{\frac{N}{2}, p_{A/B}}(N_1^{A/B}) f(N_1^{A/B}). \quad (A9)$$

- If $f(N_1^A, N_1^B) = g(N_1^A + N_1^B)$, where $g(x)$ is a function of a single variable, then

$$\langle \langle f(N_1^A, N_1^B) \rangle \rangle = \sum_{N_1^A=0}^{N/2} \sum_{N_1^B=0}^{N/2} B_{\frac{N}{2}, p_A}(N_1^A) B_{\frac{N}{2}, p_B}(N_1^B) g(N_1^A + N_1^B) \quad (A10)$$

$$= \sum_{N_1^C=0}^N \sum_{N_1^A=0}^{N_1^C} B_{\frac{N}{2}, p_A}(N_1^A) B_{\frac{N}{2}, p_B}(N_1^C - N_1^A) g(N_1^C) \quad (A11)$$

$$= \sum_{N_1^C=0}^N P_{N,p_A,p_B}(N_1^C) g(N_1^C). \quad (A12)$$

We are now equipped to look at the average of $\Delta S_{mix}/k_B = \beta F_A(N/2, \mathbf{N}_1^A, \beta, V) + \beta F_B(N/2, \mathbf{N}_1^B, \beta, V) - \beta F_C(N, \mathbf{N}_1^C, \beta, V)$. To this end, we note that the first two free energies are functions of only one of the particle number and will therefore follow Equation (A10). This case is rather straightforward and gives an expression of the form (14):

$$\begin{aligned} \langle \langle \beta F_A(N/2, \mathbf{N}_1^A, \beta, V) + \beta F_B(N/2, \mathbf{N}_1^B, \beta, V) \rangle \rangle &= -N \ln \frac{V}{2} + \frac{N}{2} (\ln [(\Lambda_1^3)^{p_A} (\Lambda_2^3)^{1-p_A}] + \ln [(\Lambda_1^3)^{p_B} (\Lambda_2^3)^{1-p_B}]) \\ &\quad + \ln \left(\frac{N}{2}! \right)^2 - \frac{N}{2} (s(p_A) + s(p_B)) \\ &\quad + H(B_{\frac{N}{2}, p_A}, B_{\frac{N}{2}, p_A}) + H(B_{\frac{N}{2}, p_B}, B_{\frac{N}{2}, p_B}). \end{aligned} \quad (A13)$$

Next, the average of the free energy once mixing has occurred follows the case described in Equation (A12) and we get:

$$\begin{aligned} \langle \langle \beta F_A(N, N_1^C, \beta, V) \rangle \rangle &= -N \ln V + \langle \langle N_1^C \rangle \rangle \ln \Lambda_1^3 + \langle \langle N - N_1^C \rangle \rangle \ln \Lambda_2^3 + \ln N! \\ &\quad - \langle \langle N_1^C \rangle \rangle \ln p_C - \langle \langle N - N_1^C \rangle \rangle \ln(1 - p_C) + H(P_{N,p_A,p_B}, B_{N,p_C}). \end{aligned} \quad (\text{A14})$$

Substituting $\langle \langle N_1^C \rangle \rangle$ by Np_C and subtracting Equation (A14) from (A13) then gives Equation (21).

References

1. Darrigol, O. The Gibbs paradox: Early history and solutions. *Entropy* **2018**, *20*, 443. [\[CrossRef\]](#)
2. Gibbs, J.W. On the Equilibrium of Heterogenous Substances. *Conn. Acad. Sci.* **1876**, *3*, 108–248.
3. Gibbs, J.W. *Elementary Principles in Statistical Mechanics*; Ox Bow Press: Woodbridge, CT, USA, 1981.
4. Glazer, M.; Wark, J. *Statistical Mechanics a Survival Guide*, 1st ed.; Oxford University Press: Oxford, UK, 2001.
5. Balescu, R. *Equilibrium and Nonequilibrium Statistical Mechanics*, 1st ed.; John Wiley & Sons: Hoboken, NJ, USA, 1975.
6. Lemons, D.S. *A student's Guide to Entropy*; Cambridge University Press: Cambridge, UK, 2013.
7. Penrose, O. *Fundation of Statistical Mechanics: A deductive Treatment*; Dover Publications: Mineola, NY, USA, 2005.
8. Balian, R. *From Microphysics to Macrophysics: Methods and Applications of Statistical Physics*; Springer: Berlin, Germany, 2003.
9. Huang, K. *Statistical Mechanics*, 2nd ed.; John Wiley & Sons: Hoboken, NJ, USA, 1987.
10. Wannier, G.H. *Statistical Physics*; Dover: Mineola, NY, USA, 1966.
11. Ben-Naim, A. *Entropy and the Second Law*; World Scientific: Singapore, 2012.
12. Allahverdyan, A.; Nieuwenhuizen, T.M. Explanation of the Gibbs paradox within the framework of quantum thermodynamics. *Phys. Rev. E* **2006**, *73*, 066119. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Unnikrishnan, C.S. The Gibbs paradox and the physical criteria for indistinguishability of identical particles. *Int. J. Quantum Inf.* **2016**, *14*, 1640037. [\[CrossRef\]](#)
14. Rosen, R. The Gibbs' Paradox and the Distinguishability of Physical Systems. *Philos. Sci.* **1964**, *31*, 232–236. [\[CrossRef\]](#)
15. Van Kampen, N.G. The Gibbs' paradox. In *Essays in Theoretical Physics: In Honor of Dirk ter Haar*; Parry, W., Ed.; Pergamon Press: Oxford, UK, 1984.
16. Jaynes, E.T. The Gibbs' paradox. In *Maximum Entropy and Bayesian Methods*; Smith, C., Erickson, G., Neudorfer, P., Eds.; Kluwer Academic: Dordrecht, The Netherlands, 1992; pp. 1–22.
17. Swendsen, R.H. Statistical mechanics of colloids and Boltzmann's definition of the entropy. *Am. J. Phys.* **2006**, *74*, 187–190. [\[CrossRef\]](#)
18. Enders, P. Is Classical Statistical Mechanics Self-Consistent? *Prog. Phys.* **2007**, *3*, 85.
19. Frenkel, D. Why Colloidal Systems Can be described by Statistical Mechanics: Some not very original comments on the Gibbs' paradox. *Mol. Phys.* **2014**, *112*, 2325–2329. [\[CrossRef\]](#)
20. Peters, H. Demonstration and resolution of the Gibbs paradox of the first kind. *Eur. J. Phys. E* **2014**, *35*, 015023. [\[CrossRef\]](#)
21. Cates, M.E.; Manoharan, V.N. Testing the Foundations of Classical Entropy: Colloid Experiments. *arXiv* **2015**, arXiv:1507.04030.
22. Paillusson, F. Gibbs' paradox according to Gibbs and slightly beyond. *Mol. Phys.* **2018**, *116*, 3196–3213. [\[CrossRef\]](#)
23. Saunders, S. The Gibbs Paradox. *Entropy* **2018**, *20*, 552. [\[CrossRef\]](#)
24. Dieks, D. The Gibbs Paradox and particle individuality. *Entropy* **2018**, *20*, 466. [\[CrossRef\]](#)
25. van Lith, V. The Gibbs Paradox: Lessons from thermodynamics. *Entropy* **2018**, *20*, 328. [\[CrossRef\]](#)
26. Sollich, P. Predicting phase equilibria in polydisperse systems. *J. Phys. Condens. Matter* **2002**, *14*, 79–117. [\[CrossRef\]](#)
27. Jacobs, W.; Frenkel, D. Predicting phase behaviour in multicomponent mixtures. *J. Chem. Phys.* **2013**, *139*, 024108. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Swendsen, R.H. Probability, Entropy, and Gibbs' Paradox(es). *Entropy* **2018**, *20*. [\[CrossRef\]](#)

29. Paillusson, F.; Pagonabarraga, I. On the role of compositions entropies in the statistical mechanics of polydisperse systems. *J. Stat. Mech.* **2014**, 2014, P10038. [[CrossRef](#)]
30. Cheraghchi, M. Expressions for the entropy of basic discrete distributions. *IEEE Trans. Inf. Theory* **2019**. [[CrossRef](#)]
31. Endres, D.M.; Schindelin, J.E. A new metric for probability distributions. *IEEE Trans. Inf. Theory* **2003**, 49, 1858–1860. [[CrossRef](#)]
32. Knessl, C. Integral representations and asymptotic expansions for Shannon and Renyi entropies. *Appl. Math. Lett.* **1998**, 11, 69–74. [[CrossRef](#)]
33. Sollich, P.; Warren, P.B.; Cates, M.E. Moment Free Energies for Polydisperse Systems. In *Advances in Chemical Physics*; John Wiley & Sons, Ltd: Hoboken, NJ, USA, 2007; pp. 265–336.
34. Enders, P. Equality and identity and (in)distinguishability in classical and quantum mechanics from the point of view of Newton's notion of state. In Proceedings of the 2008 IEEE Information Theory Workshop (ITW' 08), Porto, Portugal, 5–9 May 2008; pp. 303–307.
35. Szilard, L. On the decrease of entropy in a thermodynamic system by the intervention of intelligent beings. *Z. Phys.* **1929**, 53, 840–856. [[CrossRef](#)]
36. Sandford, C.; Seeto, D.; Grosberg, A.Y. Active sorting of particles as an illustration of the Gibbs mixing paradox. *arXiv* **2017**, arXiv:1705.05537v2.
37. Blavatska, V. Equivalence of quenched and annealed averaging in models of disordered polymers. *J. Phys. Condens. Matter* **2013**, 25, 505101. [[CrossRef](#)] [[PubMed](#)]
38. Jensen, J.L.W.V. Sur les fonctions convexes et les inégalités entre les valeurs moyennes. *Acta Math.* **1906**, 30, 175–193. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).